Analyzing treatment effects on distributions with complex structure

Mark Bebbington *Massey University*

Marcel Voia *Carleton University*

Ricardas Zitikis *University of Western Ontario*

Outline

- The Pennsylvania Bonus Experiment
- Hypotheses, and nonparametric tests
- Power, or the lack of it
- Parametric model
- More power

(apologies to Tim Taylor and Al Borland)

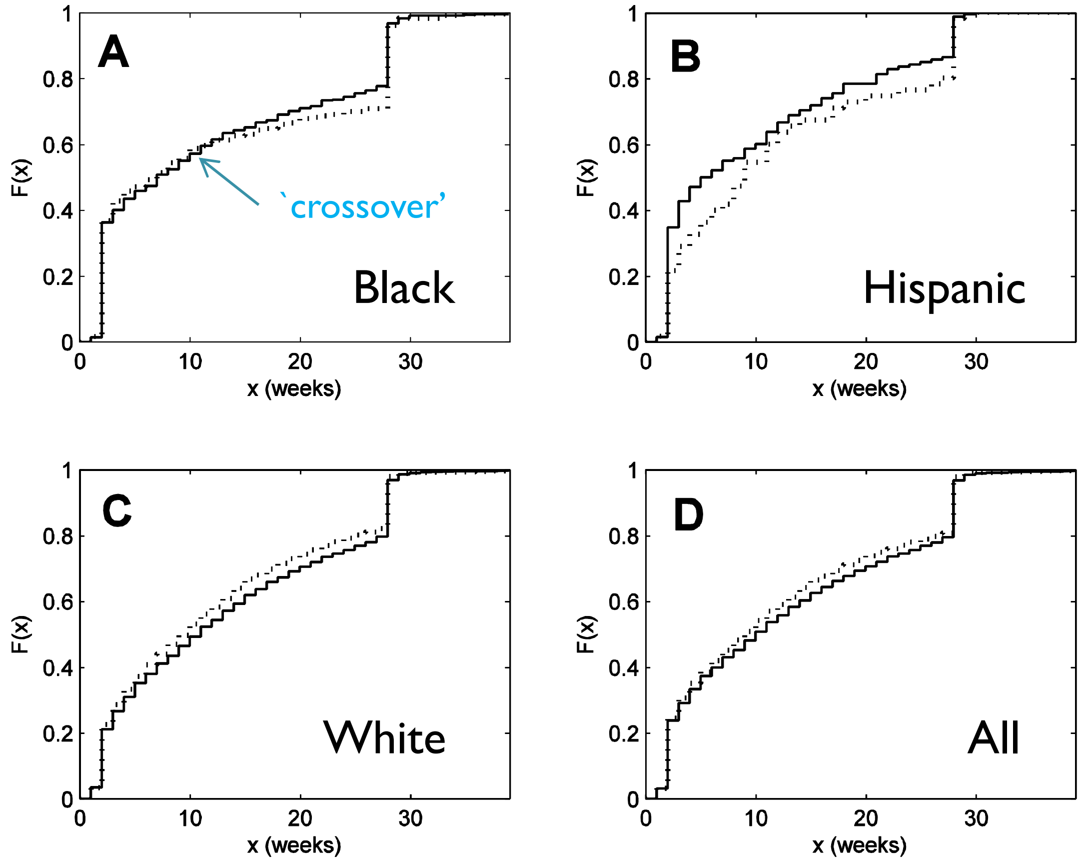
Motivation – The Pennsylvania Bonus Experiment (PBE)

- July 1988 – October 1989
- Control and Treatment groups consisting of unemployed
- Treatment comprises
  - a cash bonus (av. $997) for those reemployed within 12 weeks
  - a workshop on job hunting
- Individuals identified as 'White', 'Black' or 'Hispanic'

Data

Duration of unemployment : Treatment (dotted) and Control (solid)

Unemployment Benefit for maximum 26 weeks.



Sample sizes:

	Control (n)	Treatment (m)
Black	457	233
Hispanic	138	55
White	2979	1571
Other	21	18

Hypotheses

(Yes, I know I'm framing these *after* looking at the data …)

- For 'White', unemployment duration appears to have decreased across the time axis: $F_T(x) < F_C(x)$ for all x.
- For 'Hispanic' the effect seems to be the opposite!!
- For 'Black', something more complex is going on: $F_T(x) > F_C(x)$ for small x, $F_T(x) < F_C(x)$ for large x. – 'Crossover'

Notation and Tests

$\delta^- = \sup_x \{F_C(x) - F_T(x)\}$, $\delta^+ = \sup_x \{F_T(x) - F_C(x)\}$,
$\kappa = \max\{\delta^-, \delta^+\}$, $\tau = \min\{\delta^-, \delta^+\}$,

Test 1 $H_0 : F_C = F_T \ (\kappa = 0)$, $H_1 : F_C = F_T \ (\kappa \neq 0)$

Test 2 $H_0 : F_C \leq F_T \ (\delta^- = 0)$, $H_1 : F_C(x) > F_T(x)$ for some x $(\delta^- > 0)$
Test 3 $H_0 : F_C \geq F_T \ (\delta^+ = 0)$, $H_1 : F_C(x) < F_T(x)$ for some x $(\delta^+ > 0)$

Test 4 $H_0 : F_C \leq F_T$ or $F_C \geq F_T \ (\tau = 0)$, $H_1 : F_C >< F_T \ (\tau \neq 0)$

Let B_i indicate a Brownian bridge on [0,1], and Γ the Gaussian process defined by $\Gamma(x) = \sqrt{\eta} B_1(F_C(x)) - \sqrt{1-\eta} B_2(F_T(x))$
where η is the (asymptotic) proportion of individuals assigned to the treatment.

Nonparametrics

- Estimate all quantities by their empirical values.

THEOREM: When $\kappa = 0$, as $n,m \to \infty$,

$$P\left(\sqrt{\frac{nm}{n+m}}\hat{\kappa} > x\right) = P(\sup_z |\Gamma(z)| > x) \leq P(\sup_t |B(t)| > x) =$$

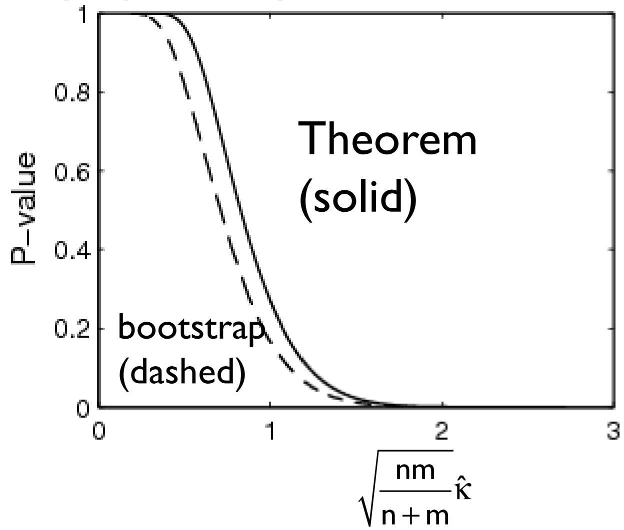
$$1 - 2\sum_{j=1}^{\infty}(-1)^{j-1}\exp\left(-2j^2x^2\right)$$

with equality if the 2 cdfs are continuous.

Results (P-Values):

	Test 1	Bootstrap
Black	0.382	0.233
Hispanic	0.390	0.236
White	0.012	0.007
All	0.064	0.035

Asymptotically …



Tests 2 and 3

- A similar theorem results, but I'll let you off this time …

P-values:

	$H_0 : F_C \leq F_T$		$H_0 : F_C \geq F_T$	
	Test 2	Bootstrap	Test 3	Bootstrap
Black	0.192	0.121	0.774	0.527
Hispanic	0.196	0.119	0.995	0.914
White	0.999	0.972	0.006	0.004
All	0.997	0.957	0.032	0.018

Test 4: The need for a parametric model…

Theorem: When $\tau = 0$, as $n,m \longrightarrow \infty$,

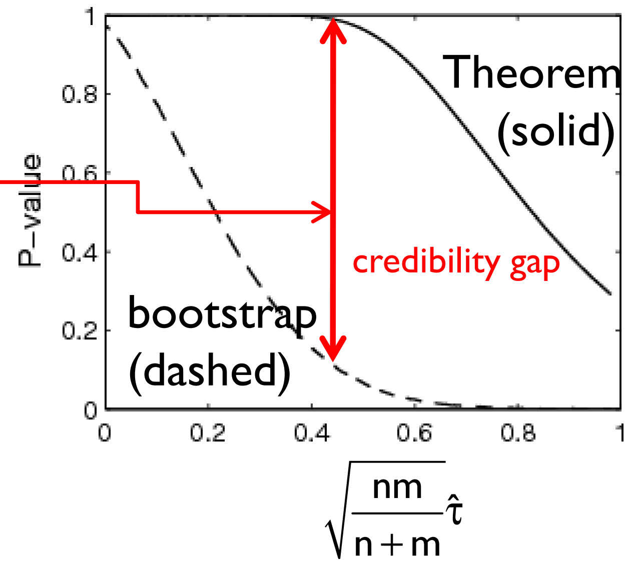
$P\left(\sqrt{\frac{nm}{n+m}}\hat{\tau} > x\right) \leq P(\sup_z |\Gamma(z)| > x)$ etc…

same bound as Test 1

The test has asymptotic power 1 if $\tau \neq 0$,

but the P-values are …

	Test 4	Bootstrap
Black	0.999	0.123
Hispanic	1	0.828
White	1	0.948
All	1	0.918



Parametric model

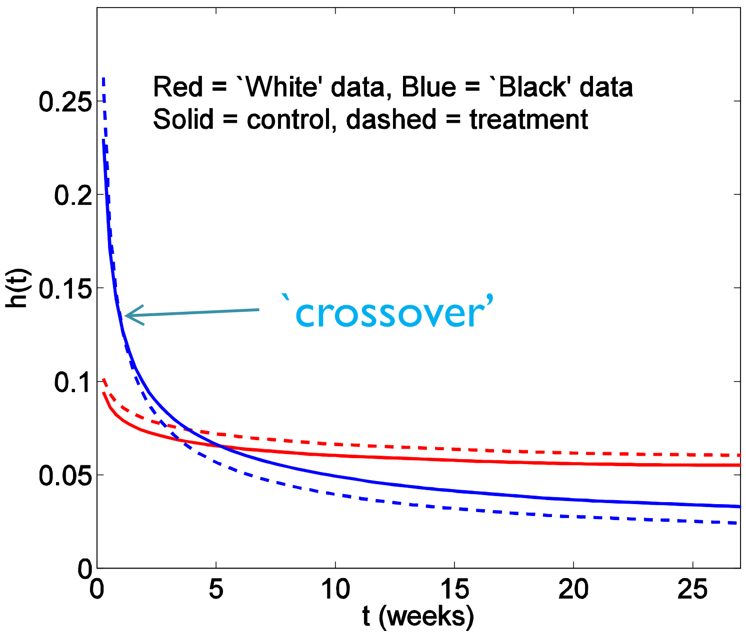
- Although data is discrete, underlying distribution (unemployment duration) is continuous.
- Discontinuity at 27 weeks (incl. 1 week stand-down)

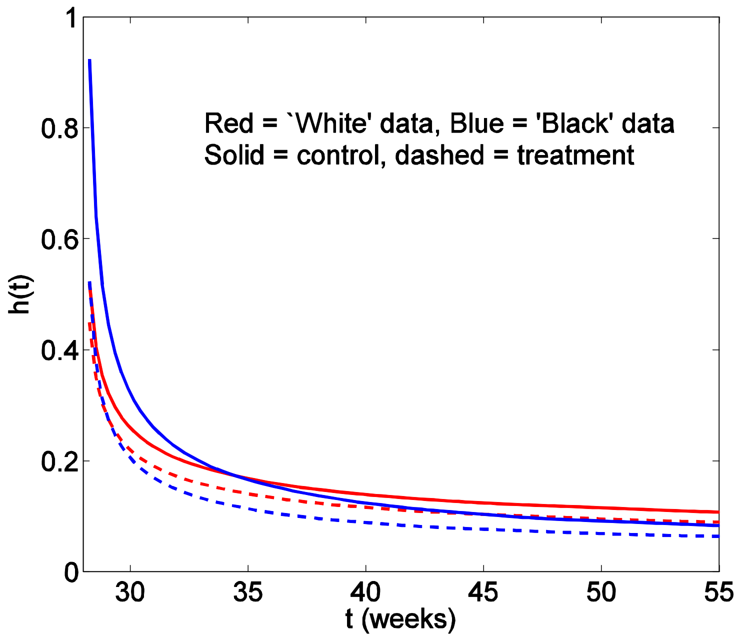
$$F(t) = \begin{cases} F_1(t) & t \leq 27 \\ F_1(27) + p(t-27) & 27 < t \leq 28 \\ F_1(27) + p + (1 - F_1(27) - p)F_2(t-28) & 28 < t \end{cases}$$

where $F_i(t) = 1 - \exp\left(-(b_i t)^{a_i}\right)$ is a Weibull distribution, allowing for inc., dec. or const. 'hazard' $ab(bt)^{a-1}$

Parameter estimates

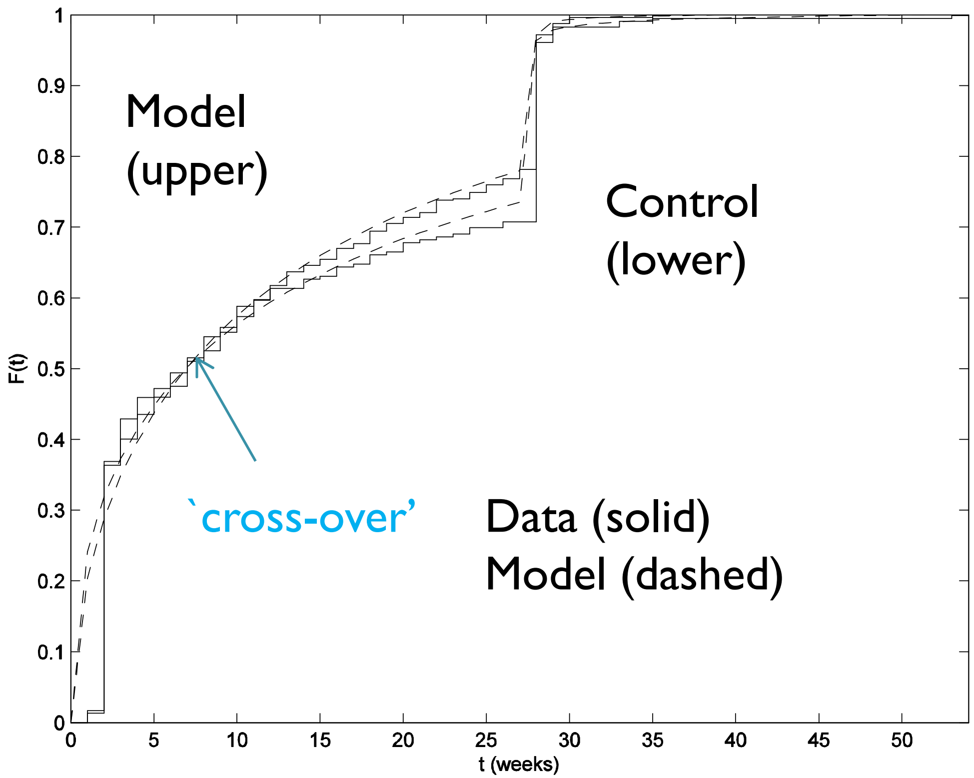
	White		Black	
	Control	Treatment	Control	Treatment
a_1	0.88	0.88	0.57	0.48
b_1	0.06	0.07	0.08	0.07
a_2	0.65	0.64	0.47	0.53
b_2	0.35	0.28	0.95	0.31
p	0.17	0.15	0.19	0.23





Power of Tests

Simulate data from model fitted to aggregated Black data (all H_0 false)



(n = N, m = N/2)

N	Test 1	Test 2	Test 3	Test 4
100	0.033	0.041	0.028	0
200	0.057	0.083	0.044	0
500	0.108	0.119	0.081	0.001
1000	0.237	0.228	0.148	0.005
2000	0.556	0.527	0.292	0.043
5000	0.988	0.934	0.794	0.584

Summary

- Asymptotic nonparametric tests
  - Provide means of identifying significant stochastic dominance (or absence thereof)
  - The test for distributional cross-over needs LARGE samples, due to the weak bound on the distribution of the test statistic
- Parametric survival analysis provides an alternative perspective on the data.
  - The PBE effect appears to be real, but is a complex one with effects from both ‘race’ and the interaction of the qualifying and benefit-eligibility periods

References

- Abadie A (2002) Bootstrap tests for distributional effects in instrumental variable models. JASA 97:284-292.
- Barret GF, Donald SG (2003) Consistent tests for stochastic dominance. Econometrica 71:71-104.
- Bilias Y (2000) Sequential testing of duration data: The case of the Pennsylvania `Reemployment Bonus' experiment. J Appl Econometrics 15:575-594.
- Horvarth L, Kokoszka P, Zitikis R (2006) Testing for stochastic dominance using the weighted McFadden-type statistic. J Econometrics 133:191-205.
- Imbens GW, Wooldridge JM (2009) Recent developments in the econometrics of program evaluation. J Econ Literature 47:5-86.
- Linton O, Maasoumi E, Whang YJ (2005) Consistent testing for stochastic dominance under general sampling schemes. Rev Econ Studies 72:735-765.

And, just for the heck of it …

- Bebbington MS, Lai CD (1996) On nonhomogeneous models for volcanic eruptions. Math Geol 28:585-600